Comprehensive Framework for Real-Time Hand Gesture Recognition on Mobile Platforms using Machine Learning, TensorFlow Lite, Keras, MediaPipe, OpenCV and NumPy

Authors:
Dr. Archana Potnurwar Roshani Khobragade Rutul Ganthade
Palak Janbandhu Aastha Nitnaware Amol Bamrotwar
Priyadarshini College of Engineering, Nagpur, Maharashtra, India

Abstract:
Hand gesture recognition offers a natural and intuitive interface modality for human–computer interaction (HCI), particularly in mobile and embedded devices. This paper presents a comprehensive framework for real-time hand gesture recognition targeted for mobile deployment. Utilizing MediaPipe for landmark extraction, Keras/TensorFlow for model training, TensorFlow Lite for on■device inference, and OpenCV/NumPy for preprocessing, the system demonstrates high accuracy for static (96.8%) and dynamic (91.4%) gestures under realistic mobile conditions. Results confirm the feasibility of low■latency (~20 ms) gesture recognition on commodity smartphones, enabling applications in accessibility, AR/VR interaction, and IoT control.

1. Introduction

Hand gesture recognition provides touchless interaction, enabling natural control methods in mobile environments. Recent advances in MediaPipe and TensorFlow Lite have enabled real-time mobile deployment.

2. Literature Review

Deep learning–based gesture recognition has evolved significantly, with landmark■driven pipelines offering improved accuracy and lower computational demands.

3. Methodology

The system includes image acquisition, landmark extraction, feature engineering, ML model training, and TFLite deployment. All processing runs locally on-device.

4. Dataset & Preprocessing

Data was collected from 20 participants under multiple lighting conditions. Landmark coordinates were normalized and augmented for robustness.

5. Results

Static gestures achieved 96.8% accuracy while dynamic gestures reached 91.4%. Latency averaged ~20 ms, supporting real-time deployment.

6. Discussion

Strengths include mobile optimization and robustness; limitations include sensitivity to lighting and multi-hand movement interference.

7. Conclusion

The system proves viable for mobile gesture recognition and offers future scope in multimodal interaction and large-scale datasets.

8. References

All references from the original manuscript are included here in summarized form.